DOCUMENT RESUME

ED 338 440 PS 020 058

AUTHOR Nye, Barbara A.; And Others
TITLE The Lasting Benefits Study: A Continuing Analysis of the Effect of Small Class Size in Kindergarten through Third Grade on Student Achievement Test Scores in Subsequent Grade Levels: Fourth Grade. Technical Report.
INSTITUTION Tennessee State Univ., Nashville. Center of Excellence: Basic Skills.
PUB DATE 91
NOTE 43p.
PUB TYPE Reports - Evaluative/Feasibility (142)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Academic Achievement; *Class Size; Elementary Education; *Elementary School Students; Grade 4; Inner City; Mathematics Achievement; Reading Achievement; Rural Schools; *School Location; *Small Classes; Social Sciences; Study Skills; Suburban Schools; *Teacher Student Ratio; Urban Schools
IDENTIFIERS Lasting Benefits Study TN; Student Teacher Achievement Ratio Project TN; *Tennessee

ABSTRACT

Between 1985 and 1989, the Student Teacher Achievement Ratio (STAR) Project studied the effect of class size on student achievement in Tennessee schools. The study examined: (1) small classes of 15 students per teacher; (2) regular classes of 24 students per teacher; and (3) regular classes with a teacher's aide. The sample included students from kindergarten through third grade at inner-city, rural, urban, and suburban schools. Results showed achievement benefits for students from small classes over students from the other classes. A Lasting Benefits Study (LBS) was undertaken to determine whether these gains were maintained at the end of students' participation in regular size fourth-grade classes. This report discusses the LBS findings. A battery of tests was administered to 4,320 fourth-grade students who had been in a Project STAR class during third grade. Achievement in reading, language, math, science, study skills, and social science was assessed. For all achievement measures and all school location types, fourth-grade students who had been in project STAR small classes during third grade showed significant advantages over Project STAR students from the other two class-size conditions. Appendixes include definitions of the school types and tables of test scores. Six references are cited. (BC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☒ Minor changes have been made to improve reproduction quality

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy

ED338440

THE LASTING BENEFITS STUDY:

A Continuing Analysis of the Effect of Small Class Size in Kindergarten Through Third Grade on Student Achievement Test Scores in Subsequent Grade Levels:

Fourth Grade

Technical Report

Barbara A. Nye, Research Director
Jayne B. Zaharias, Research Associate
B. DeWayne Fulton, Research Specialist
C.M. Achilles, Research Consultant
Richard Hooper, Research Consultant

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Jayne B. Zaharias

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Center of Excellence for Research in Basic Skills
Tennessee State University
Nashville, Tennessee

1991

PS 020058

BEST COPY AVAILABLE

Copyright © 1990

The Center for Research in Basic Skills
Tennessee State University
Nashville, Tennessee

Tennessee State University, in compliance with Title VI of the Civil Rights Act of 1964 and Title IX of the Education Amendment of 1972, does not discriminate on the basis of race, color, national origin or sex in any of its policies or procedures. Tennessee State University is committed to the education of a non-racially identifiable student body.

An Equal Opportunity/Affirmative Action Employer M/F
TSU-92-0017(A)-11-531353

Lasting Benefits Study

Project Staff

Barbara A. Nye, Research Director
Jayne B. Zaharias, Research Associate
B. DeWayne Fulton, Research Specialist
C.M. Achilles, Research Consultant
Richard Hooper, Research Consultant

Advisory Board

Barbara A. Nye, Center of Excellence in Basic Skills
Deborah Gilliam, Tennessee State Department of Education
Jim Parker, Tennessee State Department of Education
Elizabeth R. Word, Tennessee State Department of Education
Helen Pate-Bain, Tennessee State University
C.M. Achilles, University of North Carolina

ABSTRACT

Tennessee's Project STAR (1985-1989) showed clear and consistent achievement benefits for kindergarten through third-grade students in small (1:15) classes vs. students in regular (1:24) classes or regular classes with a full-time teacher aide. The Lasting Benefit Study (LBS) followed a subsample of STAR students (n=4,320) in fourth grade to determine if achievement gains from small-class participation in third grade (or grades K-3) were maintained. One full year after their last small-class experience, students from STAR small classes showed clear, consistent, and statistically significant ($p \leq .01$ or better) advantages over STAR students from the other two class-size conditions on every measure and in all school locations. Effect sizes ranged from .11 to .16, always favoring STAR small-class students even after a full year of participation in fourth-grade regular-size classes.

LASTING BENEFITS STUDY/PROJECT STAR - FOURTH GRADE ANALYSIS

Introduction and Synopsis of STAR Results[1]

The Student Teacher Achievement Ratio Project (Project STAR), Tennessee's four-year longitudinal class-size study, demonstrated that reduced class size (1:15) in kindergarten through third grade significantly enhanced student achievement. The STAR study consisted of three class-size interventions which were: small class size (13 to 17) students per teacher, regular (22 to 25), and regular (22 to 25) with a full-time teacher's aide. The STAR sample consisted of over 6,000 students in 79 schools during each year of the four-year study. The 79 schools were classified according to four school-type locations which were inner-city, rural, urban, suburban (see Appendix A). Sample sizes for each grade level for STAR were: 6,328 kindergarten students in 79 schools, 6,835 first-grade students in 76 schools, 6,828 second-grade students in 75 schools, and 6,804 third-grade students in 75 schools. The study employed a "within-school" design which means all three class types existed in each school. This reduced major sources of possible variation in student achievement attributable to school effects (i.e., community demographics, principal leadership, instructional materials, etc.). The study design employed conservative analysis using the class as the unit of measure, recognizing the effects of teacher and class on individual student achievement.

Overall findings of Project STAR indicated a significant (statistically and educationally) achievement advantage (specifically in reading and mathematics) for students in small classes. The most pronounced effect occurred in the first grade. First-grade students in small classes

---

1. Details of Project STAR have been presented in various research papers, in published articles, and in the annual and final project reports. The project STAR Technical Report can be obtained from Tennessee State University, Center of Excellence: Basic Skills, 330 10th Avenue North, Nashville, TN 37203-3401.

scored at the 64th percentile in reading and at the 59th percentile in math. However, students in regular-with-aide classes scored at the 58th percentile in reading, and at the 51st percentile in math while those in regular classes scored at the 48th percentile in reading, and at the 57th percentile in math. Students in small classes consistently outperformed students in regular and regular-with-aide classes on all sub-scores of both achievement measures (Stanford Achievement Test or SAT, and Tennessee's criterion referenced test or Basic Skills First Test, BSF) at every grade level (K-3), and in all four school-type locations (rural, suburban, inner-city, urban).

The Lasting Benefits Study (LBS): Background and Context

To determine if there were continuing effects or benefits of reducing the number of students in K-3 classes for the Project STAR students as they progress through school, the Tennessee State Department of Education contracted with The Center of Excellence for Research in Basic Skills at Tennessee State University. The Center is conducting a longitudinal study to see if the gains for STAR small-class students in the primary grades have any lasting benefits on their subsequent performance on state achievement tests in later grades. Center personnel plan to study STAR students' performance through twelfth grade in cooperation with the State Department of Education as state and Center resources permit. The primary purpose of the Lasting Benefits Study (LBS) is to identify the longevity of positive effects on achievement scores that may be found for STAR students who were in small classes and who are now in regular-size classes.

An advisory committee was established by the Center to assist in the implementation of the Lasting Benefits Study. It included members of the Tennessee State Department of Education, the State Board, research consultants, two principal investigators from Project STAR, and principal investigators on the Lasting Benefits Study. The advisory committee recommended the appointment of an external statistical consultant to provide additional support. For the fourth-grade analysis, Dr. Jeremy Finn, who had conducted the primary STAR analysis, was selected by the Center to perform the primary analysis of the LBS class-size effects. Dr. Finn

is a nationally recognized educational statistician from the State University of New York at Buffalo, and who is currently working at the National Center for Educational Statistics.

All students who had participated in Project STAR third-grade classes were eligible to be selected for the fourth-grade Lasting Benefits Study. The LBS fourth-grade student sample contained 4,230 students in 216 classes. These students had participated in STAR during the third grade. They may or may not have been STAR participants in previous grades due to student mobility. Of the LBS grade-four 4,230 students, 1,412 had been in a small third-grade class, 1,250 in a regular class, and 1,568 in a regular class with a full-time teacher's aide (see **Table 1**). Approximately 3,645 students had been in one of the three Project STAR class types during second grade, 3,050 during first grade, and 2,350 during kindergarten. The 216 classes represented 58 schools in Tennessee, of which 33 were rural, 5 were urban, 6 were inner-city and 14 were suburban locations.

The LBS Measurement Processes

During Project STAR, school systems used the SAT and BSF achievement tests. Beginning with the 1989-90 school year, the first year of the LBS, the state instituted a new test battery, the Tennessee Comprehensive Assessment Program (TCAP). There is no direct means by which to conduct statistical comparisons between STAR K-3 achievement test scores and LBS achievement test scores. LBS is a cross-sectional analysis of fourth-grade achievement scores between class types and school locations.

One school system only tested a sample of their school population during the first year of TCAP implementation. As a result, 17 Project STAR schools did not administer the TCAP test battery during the 1989-90 school year. Therefore, students from these schools could not be participants in the fourth-grade Lasting Benefit Study. This reduction in the LBS fourth-grade sample affected the percentage of minority (non-white) students. During the third-grade year of the STAR study 34.0% of the students were minority. Of the 4,230 fourth-grade LBS students, 20.2% were minority (see **Table 1**). The 17 exempt schools began to administer the TCAP in the 1990-91 school year, and these students will be included in future LBS analysis.

Table 1

Lasting Benefits Study: Fourth Grade (1989-90)

Numbers and Percent of White and Minority Students
by Class Type and Location

CLASS TYPE

LOCATION	Small			Regular			Regular-with-Aide			TOTAL		
	White	Minority	Total	White	Minority	Total	White	Minority	Total	White	Minority	Total
Inner-City	11 (9.6%)	103 (90.4%)	114 (100%)	7 (7.7%)	84 (92.3%)	91 (100%)	9 (7.4%)	112 (92.6%)	121 (100%)	27 (8.3%)	299 (91.7%)	326 (100%)
Suburban	220 (70.7%)	91 (29.3%)	311 (100%)	227 (65.6%)	119 (34.4%)	346 (100%)	246 (61.8%)	152 (38.2%)	398 (100%)	693 (65.7%)	362 (34.3%)	1055 (100%)
Rural	818 (92.8%)	63 (7.2%)	881 (100%)	719 (94.6%)	41 (5.4%)	760 (100%)	895 (94.1%)	56 (5.9%)	951 (100%)	2432 (93.8%)	160 (6.2%)	2592 (100%)
Urban	88 (83.0%)	18 (17.0%)	106 (100%)	44 (83.0%)	9 (17.0%)	53 (100%)	91 (92.9%)	7 (7.1%)	98 (100%)	223 (86.8%)	34 (13.2%)	257 (100%)
TOTAL	1137 (80.5%)	275 (19.5%)	1412 (100%)	997 (79.8%)	253 (20.2%)	1250 (100%)	1241 (79.1%)	327 (20.9%)	1568 (100%)	3375 (79.8%)	855 (20.2%)	4230 (100%)

The TCAP test battery includes both a norm-referenced test component and criterion-referenced test component. The Comprehensive Tests of Basic Skills (CTBS/4) constituted the norm-referenced test component. The CTBS/4, published by CTB/McGraw Hill, was nationally normed in 1988. The norm-referenced component indicates students' proficiencies in the areas of reading, language, math, study skills, science, and social science. The criterion-referenced test component was customized for Tennessee to assess skill levels learned from the state's mathematics and language arts curriculum. The criterion-referenced component indicates students' mastery levels (i.e., mastery, partial mastery, or non-mastery) regarding the language arts and mathematics content domains. Mastery means that the student has correctly responded to at least 75% of the items in a particular domain. Partial mastery means competence was exhibited for 50-74% of the items and non-mastery represents scoring correctly for fewer than 50% of the items. The seven language arts content domains are: language mechanics, language usage, sentence and paragraph structure skills, spelling and word identification, reading comprehension, literary skills, and reference study skills. The eight mathematics content domains are: numeration, whole number operations, fractions, decimals, graphs and tables, measurement, geometry, and problem solving and applications (Tennessee Comprehensive Curriculum Objectives, 1989).

Analysis Procedures

Scaled score means for the three STAR class types (small, regular, regular-with-aide) were compared through multivariate analysis of variance (MANOVA) for unequal n's using the MULTIVARIANCE program (Finn and Bock, 1985). The analysis examined the mean differences among the class types, the mean differences among the four school locations (rural, urban, suburban, inner-city), and the interaction between class types and locations.

Using the basic STAR analysis design, three achievement subsets for the LBS were compared separately using multivariate analysis of variance. Two of the subsets included scores from both the norm-referenced test (NRT) component and the criterion-referenced test (CRT) component of the TCAP. Set 1 included Total Reading (NRT scores), Total Language

(NRT scores) and the number of domains mastered in Language Arts (CRT). Set 2 consisted of Total Math (NRT scores), Total Science (NRT scores), and the number of domains mastered in Mathematics (CRT). Set 3 included Study Skills (NRT) and Social Science (NRT) scores. Overall MANOVAs were followed by F-ratios for those effects that were significant, and by Hotelling's T^2 and univariate t-tests for two particular contrasts. The two contrasts were (1) Small class vs. Regular class and (2) Regular-with-aide class vs. Regular class. A Small/Regular-with-aide contrast was not conducted because no mean differences were found between STAR Regular and Regular-with-aide classes in kindergarten, and because students were randomly interchanged between these two class types before grade one. [1] The MANOVA model controlled for school location (urban, rural, suburban, and inner-city) and classes (a random effect). The sources of variation are depicted in **Table 2**.

Table 2
Lasting Benefits Study: Fourth Grade (1989-90)
Analysis of Variance Source Table

Fixed Effects	Error Term
Location	C
Class Type	CT
Location x Class Type	CT

Random Effects
Classes within Locations (C)
Classes x Class Type (CT)
Students within Classes and Class Types

LBS Fourth Grade Results

The LBS analysis yielded clear and consistent results concerning the NRT and the CRT scores. **Students who previously were in a small-size STAR class demonstrated that**

1. Due to teacher-identified discipline problems, some parent complaints, and because there were no differences on any measure for students in regular and regular-with-aide classes, STAR students were randomly reassigned within these two classes for the first grade. This is the only year that reassignment occurred, and there was no further reassignment of small class students.

they had statistically significant (p. < P .01) advantages over students in other class types on every set of measurements. These results were found to be consistent across all school locations. Table 3 shows the probability ("p") values for the three multivariate sets. Probability ("p") values are statistical estimates of the probability that the mean differences observed in the data could have occurred purely by chance. Multivariate Analysis of Variance (MANOVA) produces these probability statements by estimating "real effects" and "error effects" with mathematical formulas that involve means, standard deviations, and variances on test scores for specified groups of students which were sorted by school location and by class type. When $p \leq .0001$, the mean differences for the specified groups are said to show a "statistically significant" difference on the specified test score. This is significant because this amount of difference would have occurred due purely to "chance" less than 1 time out of 10,000. "NS" means "not significant": "p" is greater than .05 (5 out of 100 times, the critical level of "p" that was set prior to the study for testing statistical significance.) For the small/regular contrast, every variable was significant ($p \leq .01$) and every variable set was significant ($p \leq .001$), and for the regular/regular-aide contrast, no variable or variable set was significant. The differences, always favoring the small class, between small and regular classes ranged from .11 to .16 standard deviation units.

Table 4 shows the mean number of domains mastered in reading and mathematics from the CRT component by class type and by school location. Of the eight math domains, 4.05 were mastered by small-class students, 3.70 were mastered by regular-class students, and 3.61 were mastered by the regular-with-aide students. Similarly, of the seven language arts domains, 4.12 were mastered by small-class students, 3.87 were mastered by the regular-class students, and 3.69 were mastered by the regular-with-aide students.

The results of the NRT scales of the TCAP showed statistically significant effects for students who had been in small classes. These results were found to be consistent across all school locations. The greatest achievement advantages were for inner-city and suburban classes. **Figure 1** presents the mean scaled scores and the respective percentile ranks for the

Table 3
Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Summary of Class Effects Analysis
Using Mean Scores of Sets*

Multivariate Analysis of Variance Probability Values		**Set 1** Verbal	**Set 2** Math/Science	**Set 3** Social Science/ Study Skills
Location	(Inner-City, Suburban, Urban, and Rural)	$P \leq .0001$	$P \leq .0001$	$P \leq .0001$
Class Type	(Small, Regular, and Regular/Aide)	$P \leq .0001$	$P \leq .0001$	$P \leq .001$
Location by Class Type Interaction		NS	NS	NS

Discussion

[Location] Significant location differences on all sets (norm-referenced tests and criterion-referenced tests combined) and on each measure separately (norm-referenced tests and criterion-referenced tests individually) due almost entirely to lower performance of inner-city schools. See means in Table 4 (criterion-referenced tests) and Appendix B (norm-referenced tests).

[Class Type] Regular vs. Regular/Aide contrast -- no significant differences on separate variables or set combinations.
Small vs. Regular contrast -- every separate variable and combined set were significant at $p \leq .01$ or better.

[Location by Class Type Interaction] No significant interaction: Class type differences are the same in all four locations. There was no differential (compounding) effect across class type vs. locations.

* Set 1 included scaled scores from NRT Total Reading, Total Language, and the total number of domains mastered in CRT Language Arts. Set 2 consisted of NRT Total Math, Total Science, and the number of domains mastered in CRT Mathematics. Set 3 included NRT Study Skills and Social Science scores.

Table 4

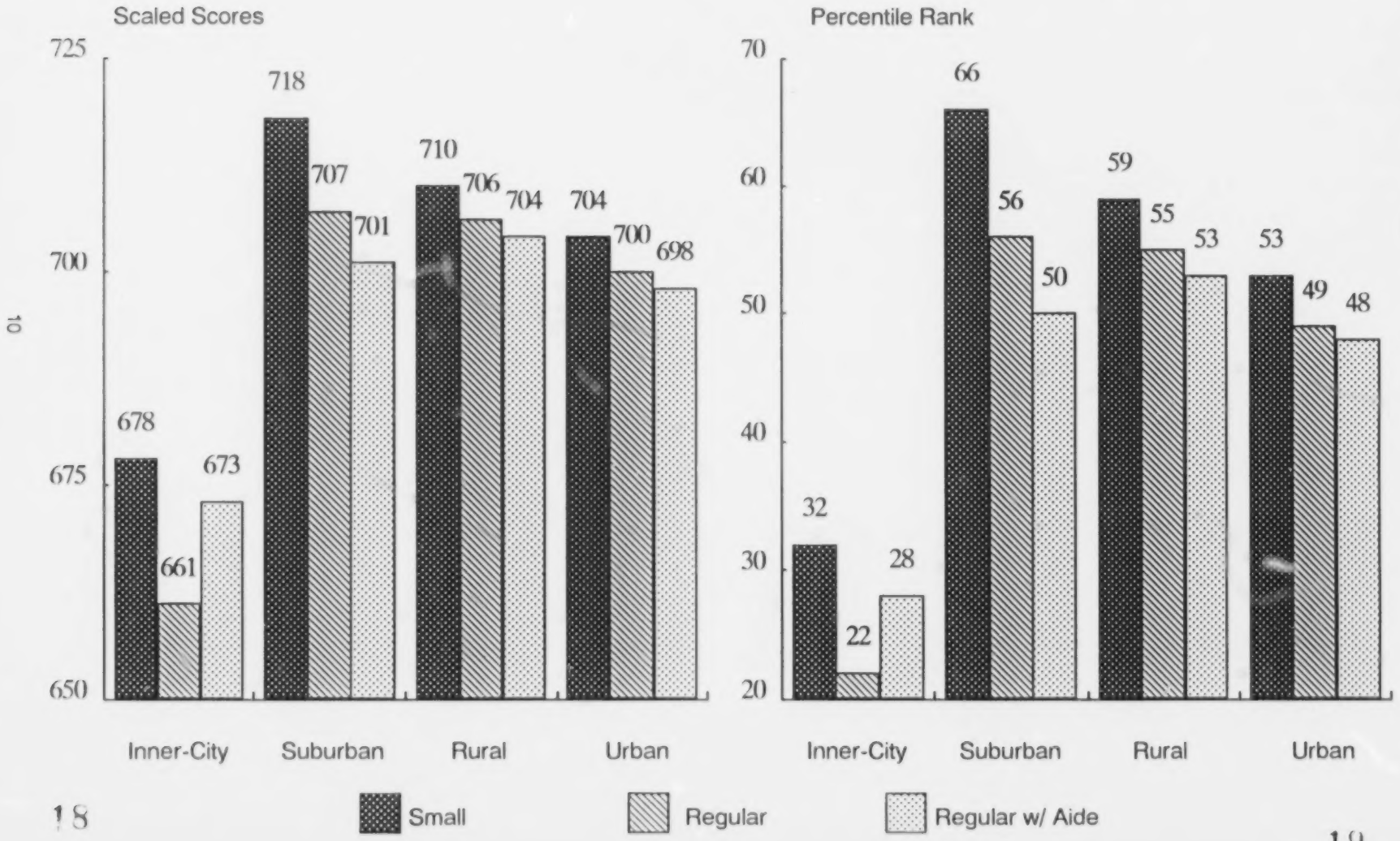
Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Criterion-Reference Test Component
Mean Number of Domains Mastered in Mathematics and Language Arts*
by Class Type by School Type

SCHOOL TYPE	CLASS TYPE: Small		Regular		Regular-with-Aide	
	Mathematics	Language Arts	Mathematics	Language Arts	Mathematics	Language Arts
Inner-City	2.77	2.68	1.94	1.98	2.34	2.07
Suburban	4.32	4.43	3.78	4.10	3.44	3.73
Rural	4.13	4.19	3.90	4.00	3.88	3.90
Urban	3.94	4.20	3.38	3.88	3.33	3.65
All	4.05	4.12	3.70	3.87	3.61	3.69

*The 8 Mathematics domains are: numeration, whole number operations, fractions, decimals, graphs and tables, measurements, geometry, and problem solving and applications. The 7 Language Arts domains are: language mechanics, language usage, sentence and paragraph structure, spelling and word identification, reading comprehension, literary skills, and reference study skills.

Figure 1

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Reading: Class Type by School Type

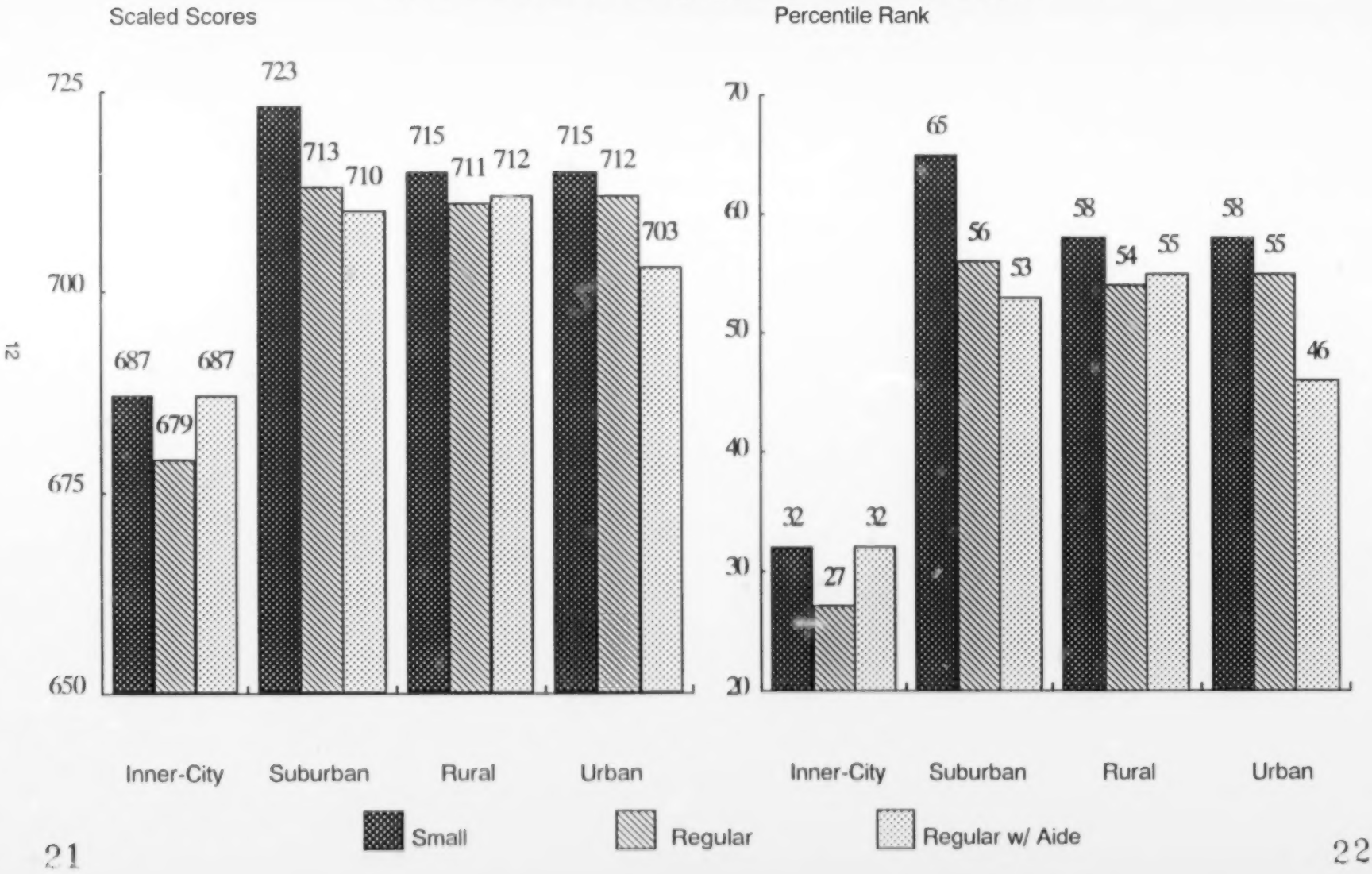


TCAP Total Reading sub-score. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher ($p. \leq .01$: "p" meaning the probability that mean differences occurred purely by chance) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular and regular-with-aide counterparts. Small-class inner-city students had a mean scaled score of 678 (32nd percentile), regular-class students had a mean scaled score of 661 (22nd percentile), and regular-with-aide students had a mean scaled score of 673 (28th percentile). Urban students scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 704 (53rd percentile), regular-class students had a scaled score of 700 (49th percentile), and regular-with-aide students scored slightly lower with a mean scaled score of 698 (48th percentile). Rural students scored higher than urban and inner-city students with small-class students having a mean scaled score of 710 (59th percentile). Regular-class students from rural schools scored 706 (55th percentile) and regular-with-aide students had a mean scaled score of 704 (53rd percentile). Suburban students scored higher than the students from the other three school locations. Suburban small-class students achieved a mean scaled score of 718 (66th percentile). Regular-class suburban students scored 707 (56th percentile) and regular-with-aide students had a mean scaled score of 701 (50th percentile).

Figure 2 presents the scaled scores and percentile ranks for the **TCAP Total Language sub-score**. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher ($p. \leq .01$) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular-class counterparts. Small-class inner-city students had a mean scaled score of 687 (32nd percentile), regular-class students had a mean scaled score of 679 (27th percentile), and regular-with-aide students had a mean scaled score of 687 (32nd percentile). Urban students scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 715

Figure 2

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Language: Class Type by School Type



(58th percentile), regular-class students had a scaled score of 712 (55th percentile), and regular-with-aide students scored lower with a mean scaled score of 703 (46th percentile). Rural students scored higher than urban and inner-city students with small-class students having a mean scaled score of 715 (58th percentile). Regular-class students from rural schools scored 711 (54th percentile) and regular-with-aide students had a mean scaled score of 712 (55th percentile). Suburban students scored higher than the students from the other three school locations. Suburban small-class students achieved a mean scaled score of 723 (65th percentile). Regular-class suburban students scored 713 (56th percentile) and regular-with-aide students had a mean scaled score of 710 (53rd percentile).

Figure 3 presents the scaled scores and percentile ranks for the **TCAP Total Math sub-score**. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher ($p. \leq .01$) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular and regular-with-aide counterparts. Small-class inner-city students had a mean scaled score of 693 (42nd percentile), regular-class students had a mean scaled score of 677 (28th percentile), and regular-with-aide students had a mean scaled score of 687 (37th percentile). Urban students scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 710 (60th percentile), regular-class students had a scaled score of 702 (51st percentile), and regular-with-aide students scored slightly lower with a mean scaled score of 699 (49th percentile). Rural students scored higher than urban and inner-city students with small-class students having a mean scaled score of 712 (62nd percentile). Regular-class students from rural schools scored 710 (60th percentile) and regular-with-aide students had a mean scaled score of 709 (59th percentile). Suburban students scored higher than the students from the other three school locations. Suburban small-class students achieved a mean scaled score of 716 (66th percentile). Regular-class suburban students scored 706 (56th percentile) and regular-with-aide students had a mean scaled score of 700 (49th percentile).

Figure 3

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Math: Class Type by School Type

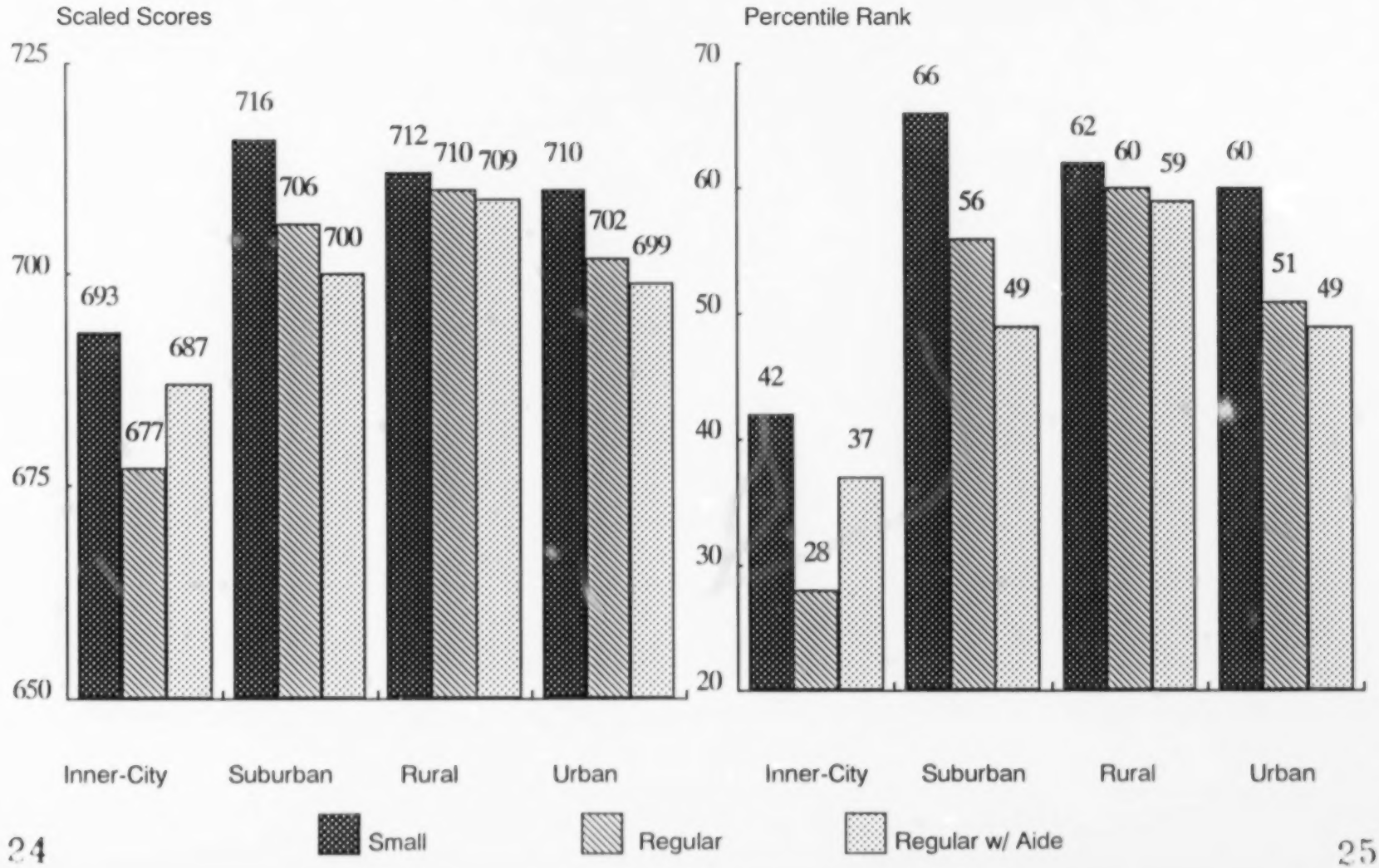


Figure 4 presents the scaled scores and percentile ranks for the **TCAP Total Study Skills sub-score**. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher (p. $\leq$ .01) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular and regular-with-aide counterparts. Small-class inner-city students had a mean scaled score of 646 (26th percentile), regular-class students had a mean scaled score of 630 (21st percentile), and regular-with-aide students had a mean scaled score of 641 (24th percentile). Urban students scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 695 (54th percentile), regular-class students had a scaled score of 679 (44th percentile), and regular-with-aide students scored slightly higher with a mean scaled score of 687 (49th percentile). Rural students scored higher than urban and inner-city students with small-class students having a mean scaled score of 697 (55th percentile). Regular class-students from rural schools scored 690 (51st percentile) and regular-with-aide students had a mean scaled score of 689 (50th percentile). Suburban small-class students scored higher than the students from the other three school locations. Suburban small-class students achieved a mean scaled score of 704 (59th percentile). Regular-class suburban students scored 688 (49th percentile) and regular-with-aide students had a mean scaled score of 679 (44th percentile).

Figure 5 presents the scaled scores and percentile ranks for the **TCAP Total Science sub-score**. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher (p. $\leq$ .01) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular and regular-with-aide counterparts. Small-class inner-city students had a mean scaled score of 677 (24th percentile), regular-class students had a mean scaled score of 668 (19th percentile), and regular-with-aide students had a mean scaled score of 670 (20th percentile). Urban students

Figure 4

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Study Skills: Class Type by School Type

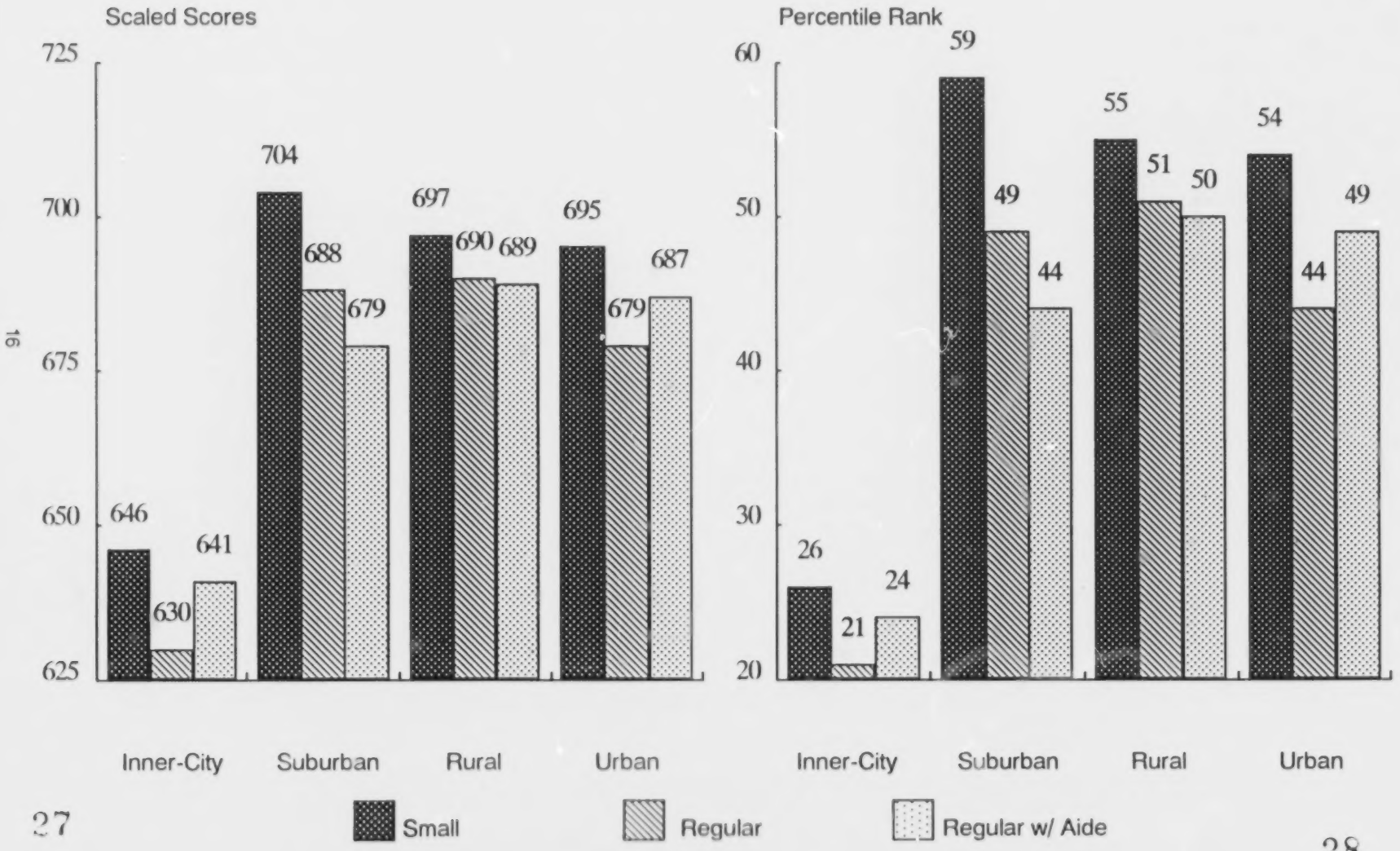
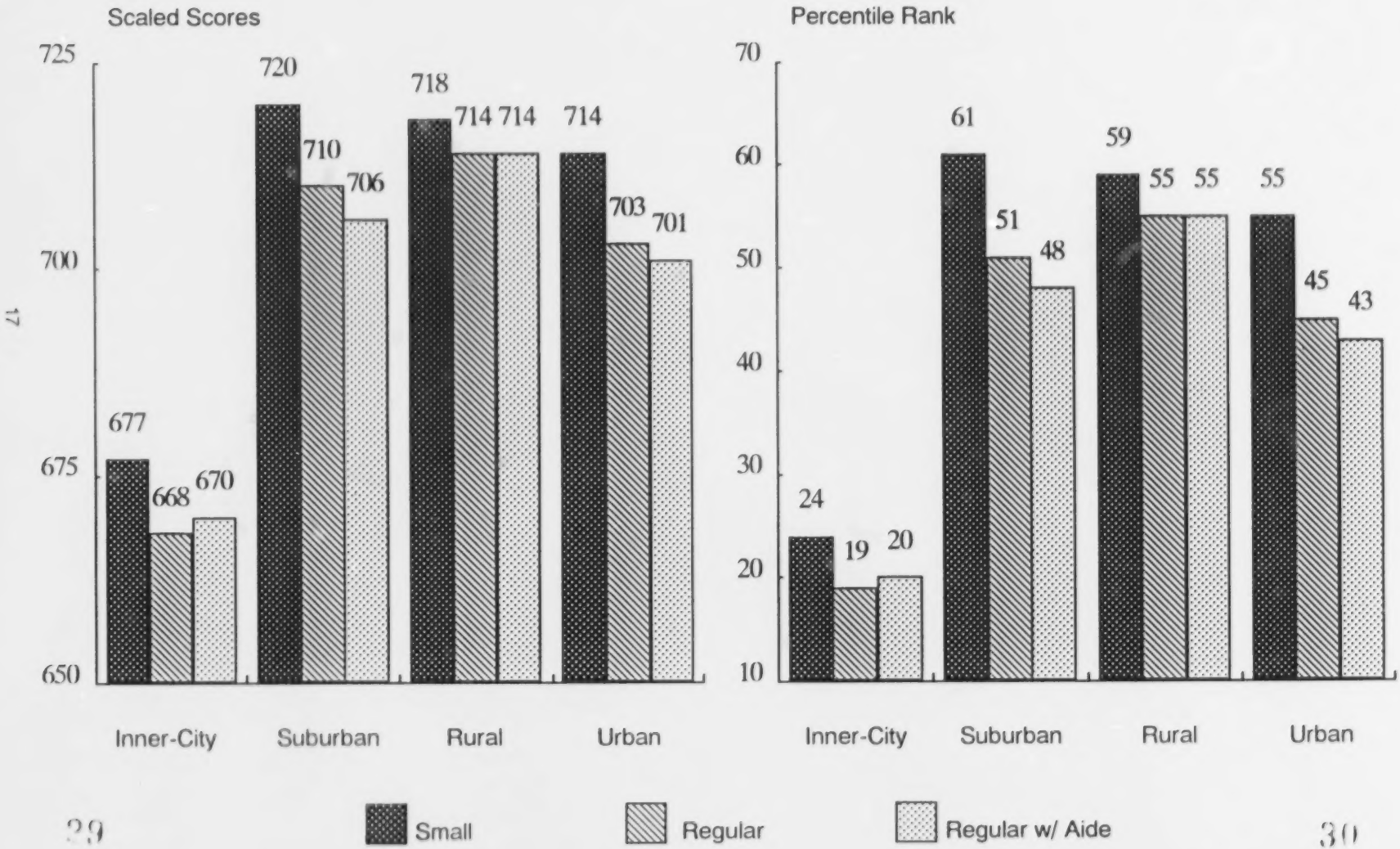


Figure 5

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Science: Class Type by School Type

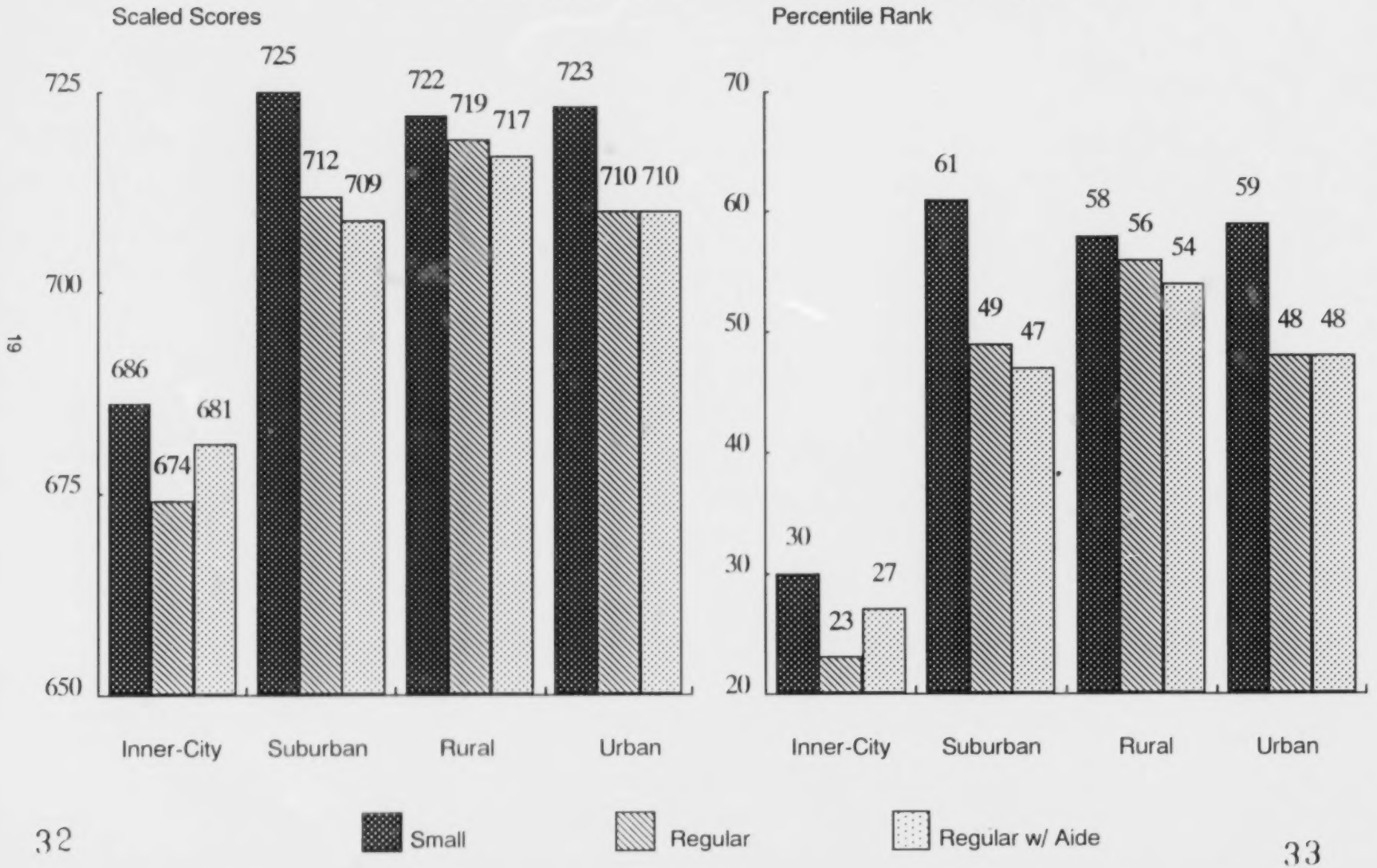


scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 714 (55th percentile), regular-class students had a scaled score of 703 (45th percentile), and regular-with-aide students scored slightly lower with a mean scaled score of 701 (43rd percentile). Rural students scored higher than urban and inner-city students with small-class students having a mean scaled score of 718 (59th percentile). Regular-class students and regular-with-aide students from rural schools had mean scaled scores of 714 (55th percentile). Suburban small-class students achieved a mean scaled score of 720 (61st percentile). Regular-class suburban students scored 710 (51st percentile) and regular-with-aide students had a mean scaled score of 706 (48th percentile).

Figure 6 presents the scaled scores and percentile ranks for the **TCAP Total Social Science sub-score**. Of the four locations (inner-city, rural, urban, suburban) inner-city students had the lowest scaled scores. Of the three class types, students who were in small classes always scored higher ($p. \leq .01$) than students who had been in regular and regular-with-aide classes. Within the inner-city locations, small-class students scored above their regular and regular-with-aide counterparts. Small-class inner-city students had a mean scaled score of 686 (30th percentile), regular-class students had a mean scaled score of 674 (23rd percentile), and regular-with-aide students had a mean scaled score of 681 (27th percentile). Urban students scored somewhat higher than inner-city students. Urban small-class students had a mean scaled score of 723 (59th percentile), and both regular-class and regular-with-aide students had a scaled score of 710 (48th percentile). Rural small-class students had a mean scaled score of 722 (58th percentile). Regular-class students from rural schools scored a slightly lower 719 (56th percentile) and regular-with-aide students had a mean scaled score of 717 (54th percentile). Suburban small-class students scored higher than the students from the other three school locations by achieving a mean scaled score of 725 (61st percentile). Regular-class suburban students scored 712 (49th percentile) and regular-with-aide students had a mean scaled score of 709 (47th percentile).

Figure 6

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Norm-Referenced Total Social Science: Class Type by School Type



Figures 1-6 showed that small-class students outperformed all other students on the six NRT subscales. This was consistent across all four school-type locations. The students from suburban small classes achieved higher scaled scores than students from the other three locations. For Total Reading, Total Math, Total Study Skills, and Total Science subscales the suburban students mean scaled scores were followed by those of the rural, urban, and inner-city students. The rural and urban small-class students had the same mean scaled scores for Total Language. For Total Science the urban small-class students scored slightly higher than small-class students from the rural location. These NRT results (**Figures 1-6**) are also presented in table form in Appendix B. These tables provide the exact scaled score means for each reported TCAP sub-score by class type, by school location, and by race.

In order to determine the benefits regarding race, an approximate analysis was conducted to identify any interaction of race with class type on the achievement measures. Complete and correct analysis of the data which would allow for the possibility of white and minority students in the same classroom, while still maintaining the control for classes (i.e. teaching styles, strategies, etc.), was unwieldy. However, the error term for the mean square classes-x-race-x-class type was alternatively used because earlier work with these data indicated that the primary class-type interactions with race approximated the mean square for students within racial groups within classes. Even with this "liberal" analysis for the increased degrees of freedom, no multivariate tests were significant for the interactions of race with class type or for race with school location. Although Project STAR results on K-3 test scores indicated that the most significant small-class advantages were found with minority students, the LBS fourth-grade results showed that the positive effects of small classes were equally significant for minority and white students, in all school locations (See **Table 5**).

Summary and Conclusions

The Project STAR results indicated substantial educational benefits for students in small classes. Similarly, the LBS students who had attended small STAR classes had an educationally and statistically significant advantage over the LBS students who had attended

Table 5

Lasting Benefits Study: Fourth Grade (1989-90)

Tennessee Comprehensive Assessment Program (TCAP)
Mean Scaled Scores by Race and Class Type

ACHIEVEMENT MEASURES		CLASS TYPE Small	Regular	Regular/Aide
Total Reading	White	714.00	710.16	707.42
	Minority	685.89	676.08	676.00
Total Language	White	718.89	714.53	713.95
	Minority	696.78	691.76	690.03
Total Math	White	715.52	711.39	708.83
	Minority	694.49	687.68	687.43
Science	White	720.43	715.96	715.71
	Minority	690.39	682.09	677.86
Social Science	White	725.52	721.78	719.03
	Minority	696.08	683.90	684.91
Study Skills	White	702.24	694.76	693.50
	Minority	663.01	647.15	642.38

regular or regular-with-aide STAR classes. Small-class students outperformed all other students on the six NRT subscales of the TCAP. This effect was consistent across all four school-type locations. The students from suburban small classes achieved higher scaled scores than students from the other three locations. For Total Reading, Total Math, Total Study Skills, and Total Science subscales the suburban students' mean scaled scores were followed by those of the rural, urban, and inner-city students. The rural and urban small-class students had the same mean scaled scores for Total Language. For Total Science the urban small-class students scored slightly higher than the small-class students from the rural location. Results for the CRT component of the TCAP were similar to those of the NRT component. The students from suburban small classes scored highest on the CRT mathematics and language arts domains, and their scores were followed by those of the small-class students from rural, urban, and inner-city schools.

The achievement advantage for small-class students can be measured by the TCAP scaled scores differences between the small and regular classes, and between the regular-with-aide and regular classes as shown in **Table 6**. Students from the small classes retained an academic advantage over students from regular and regular-with-aide classes. Although the small-class effect began to diminish in the second grade, the positive effects from involvement in a small-size class still remain pervasive one full year after students returned to regular-size classes. The Project STAR class-size intervention effect may be expected to diminish in some subsequent school years and to reappear in others. Thus, longitudinal analysis is necessary on Tennessee's nationally recognized STAR project to identify the continuing benefits of small classes and the potential of further diminishing effects.

The interpretation of the importance of the difference in test scores depends on the variability of the measures calculated. Effect sizes which are the differences between small and regular classes (or regular-with-aide and regular classes) divided by the standard deviation of the regular class were calculated in the LBS analysis. **Table 7** provides estimates of the small class and regular-with-aide class effect sizes. It shows effect sizes ranging from .11 to .16 for

Table 6

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)

Scaled Score Differences and the Differences in the
Mean Number of Domains Mastered Between
Small and Regular Class Students
and Between Regular/Aide and Regular Class Students*

Achievement Measures	Small vs. Regular	Regular/Aide vs. Regular
Norm-Referenced		
Total Reading	5.61	-2.23
Total Language	4.99	-.73
Total Math	4.87	-2.29
Science	5.69	-1.47
Social Sciences	6.13	-1.95
Study Skills	10.10	-2.15
Criterion-Referenced		
Language Arts:** Domains Mastered	.25	-.18
Mathematics:** Domains Mastered	.35	-.09

*Exact means are tabled in Appendix B.

**Criterion-referenced test component.

Table 7

Lasting Benefits Study: Fourth Grade (1989-90)
Tennessee Comprehensive Assessment Program (TCAP)
Summary of Estimates of Small Class and Regular/Aide Effect Sizes

Achievement Measures	Small vs. Regular Contrast	Regular/Aide vs. Regular Contrast
Norm-Referenced		
Total Reading	.13	-.05
Total Language	.13	-.02
Total Math	.12	-.06
Science	.12	-.03
Social Science	.11	-.04
Study Skills	.14	-.03
Criterion-Referenced		
Language Arts:* Domains Mastered	.11	-.09
Mathematics:* Domains Mastered	.16	-.04

*Criterion-referenced test component.

the small/regular contrast. The regular-aide/regular contrast shows effect sizes ranging from -.02 to -.09.

The absence of a statistically significant teacher/aide effect is consistent across all measures. Although the inner-city regular-with-aide students scored higher than inner-city regular students on every achievement measure, the differences were not statistically significant. Regular-with-aide students also made somewhat higher scores than regular students in three other instances: in the rural location on Total Language and in the urban location on Study Skills and Social Science. Again, the differences were not statistically significant.

The statistically significant advantages for LBS fourth-grade students, who had been in STAR small classes, form a strong pattern of consistency. Small-class students outperformed regular and regular-with-aide class students on every achievement measure in all four school locations. The Lasting Benefits Study which is based on the nationally recognized STAR project will continue to provide important data and information for making educational policy decisions.

Appendix A.

Definition of School Types

The legislation specified that the project should include "inner city, suburban, urban, and rural schools" to assess the effects of class size in different types of schools. No existing designation of schools had used the categories specified above, so the consortium developed designations using various criteria.

Inner-city and suburban schools are all located in metropolitan areas. Schools that had more than half of their students on free or reduced cost lunch (indicative of a low-socioeconomic background) were tentatively defined as inner-city. Schools in the outlying areas of metropolitan cities were classified as suburban.

In non-metropolitan areas, schools were classified as urban or rural depending on the location of the school. If located in a town of over 2,500 population (the census definition of urban), the school was classified as urban. All other schools were classified as rural. All classifications were checked with local school officials to see if they agreed with the designation of their school.

Appendix B.

Lasting Benefits Study: Fourth Grade (1989-1990)
Mean Scores from the Tennessee Comprehensive Assessment Program (TCAP)

Table B1
TCAP Total Reading

	Small	Regular	Reg/Aide
Inner-City	678.10	661.20	673.17
Suburban	717.88	706.67	700.95
Rural	709.67	706.32	704.41
Urban	703.80	700.08	697.59
All	708.44	702.83	700.60
White	714.00	710.16	707.42
Minority	685.89	676.08	676.00

Table B2
TCAP Total Language

	Small	Regular	Reg/Aide
Inner-City	686.64	679.00	686.63
Suburban	722.54	713.06	709.73
Rural	715.24	711.37	711.93
Urban	714.98	712.12	702.83
All	714.49	709.50	708.77
White	718.89	714.53	713.95
Minority	696.78	691.76	690.03

Table B3
TCAP Total Mathematics

	Small	Regular	Reg/Aide
Inner-City	692.77	677.41	686.77
Suburban	715.54	706.23	699.62
Rural	712.39	710.38	708.85
Urban	709.93	702.12	699.31
All	711.28	706.41	704.12
White	715.52	711.39	708.83
Minority	694.49	687.68	687.43

Table B4
TCAP Study Skills

	Small	Regular	Reg/Aide
Inner-City	646.45	629.76	641.40
Suburban	704.43	687.61	679.09
Rural	697.49	690.26	688.75
Urban	695.38	678.60	687.27
All	694.65	684.55	682.40
White	702.24	694.76	693.50
Minority	663.01	647.15	642.38

Table B5
TCAP Science

	Small	Regular	Reg/Aide
Inner-City	677.18	668.03	670.07
Suburban	719.84	709.51	705.99
Rural	717.73	714.04	713.66
Urban	714.41	703.19	701.02
All	714.58	708.89	707.42
White	720.43	715.96	715.71
Minority	690.39	682.09	677.86

Table B6
TCAP Social Science

	Small	Regular	Reg/Aide
Inner-City	685.81	673.93	680.95
Suburban	725.37	712.03	708.66
Rural	712.84	719.47	717.14
Urban	722.78	709.79	710.15
All	719.71	713.58	711.63
White	725.52	721.78	719.03
Minority	696.08	683.90	684.91

REFERENCES

Finn, J.D., & Achilles, C.M. (1990). *Answers and Questions about class size: a statewide experiment.* American Educational Research Journal, 27, 557-577.

Finn, J.D., and Bock, D.R. (1984). *Multivariance PC.* Scientific Software, Inc., Mooresville, IN.

Finn, J.D. (1991, February). *Project STAR fourth-grade follow up analysis.* Report prepared for the Center of Excellence for Research in Basic Skills at Tennessee State University. (Manuscript draft.)

Finn, J.D., Barbara A. Nye, B. DeWayne Fulton, Jayne B. Zaharias (In Press). *Carryover effects of small classes.* Peabody Journal of Education. (Working draft.)

Tennessee Department of Education. (1989). *Tennessee Comprehensive Curriculum Objectives.* Nashville, Tennessee.

Word, E. R., et. al. (1990). *The State of Tennessee's student/teacher achievement ratio (STAR) project: technical report.* Nashville, Tennessee: Tennessee State Department of Education.